

# **SRM VALLIAMMAI ENGINEERING COLLEGE**

**(An Autonomous Institution)  
SRM Nagar, Kattankulathur – 603 203**

**DEPARTMENT OF COMPUTER APPLICATIONS**

**QUESTION BANK**



**III SEMESTER M.C.A.**

**PMC305 - EXPLORATORY DATA ANALYSIS**

**Regulation – 2024**

**Academic Year 2025 – 2026 (Odd Semester)**

*Prepared by*

**Mr. M. ASAN NAINAR (AP Sel.G/M.C.A.)**

# SRM VALLIAMMAI ENGINEERING COLLEGE

(An Autonomous Institution)

SRM Nagar, Kattankulathur-603 203

DEPARTMENT OF COMPUTER APPLICATIONS

## QUESTION BANK

**SUBJECT : PMC305 - EXPLORATORY DATA ANALYSIS**

**SEM/YEAR: III / II**

### **UNIT - I: EXPLORATORY DATA ANALYSIS**

EDA fundamentals – Understanding data science - Significance of EDA – Making sense of data – Comparing EDA with classical and Bayesian analysis – Software tools for EDA – Visual Aids for EDA –Data transformation techniques-merging database, reshaping and pivoting, Transformation techniques.

### **UNIT - I: PART – A**

<b>Q. No</b>	<b>Question</b>	<b>BT Level</b>	<b>Competence</b>	<b>Course Outcome</b>
1	Define Exploratory Data Analysis.	BTL-1	Remember	CO1
2	What is the main goal of EDA in data science?	BTL-1	Remember	CO1
3	Mention two popular tools used for EDA.	BTL-1	Remember	CO1
4	What does a histogram reveal about a dataset?	BTL-1	Remember	CO1
5	State two advantages of using visual aids in EDA.	BTL-1	Remember	CO1
6	Define skewness in a dataset.	BTL-1	Remember	CO1
7	Name two Python libraries commonly used for EDA.	BTL-1	Remember	CO1
8	Define data reshaping with an example.	BTL-1	Remember	CO1
9	What is data pivoting in EDA?	BTL-1	Remember	CO1
10	Mention any two transformation techniques used in EDA.	BTL-1	Remember	CO1
11	Name any two statistical measures used in EDA.	BTL-1	Remember	CO1
12	List the benefit of using pivot tables.	BTL-1	Remember	CO1
13	What is data slicing?	BTL-1	Remember	CO1
14	How is missing data typically handled in EDA?	BTL-2	Understand	CO1
15	Differentiate between univariate and multivariate analysis.	BTL-2	Understand	CO1
16	Differentiate between EDA and confirmatory data analysis,	BTL-2	Understand	CO1
17	How does EDA contribute to model selection?	BTL-2	Understand	CO1
18	How is EDA different from classical statistical methods?	BTL-2	Understand	CO1

19	How does Bayesian analysis differ from classical statistics?	BTL-2	Understand	CO1
20	Write any two advantages of merging datasets during analysis.	BTL-2	Understand	CO1
21	Distinguish between mean and median.	BTL-2	Understand	CO1
22	Why is EDA considered an iterative process?	BTL-2	Understand	CO1
23	How do reshaping and pivoting help in better data representation?	BTL-2	Understand	CO1
24	Write the purpose of using a heatmap in data visualization.	BTL-2	Understand	CO1
25	What are categorical variables? Give one example.	BTL-1	Remember	CO1

### UNIT - I: PART – B

CO1

Q. No	Question	Mark	BT Level	Competence	Course Outcome
1	Explain different types of data transformations and their effects on modeling.	16	BTL-3	Apply	CO1
2	Explain the impact of data reshaping and pivoting on feature extraction and model building.	16	BTL-3	Apply	CO1
3	Illustrate EDA techniques applied to time-series data.	16	BTL-3	Apply	CO1
4	Discuss the limitations of EDA and how modern tools have helped overcome them.	16	BTL-3	Apply	CO1
5	(i) Describe pivoting and reshaping operations using pandas in Python. (ii) Include code examples and use cases.	10 6	BTL-3	Apply	CO1
6	Discuss the ethical considerations while performing EDA on sensitive data.	16	BTL-3	Apply	CO1
7	(i) Explain how EDA is useful in data cleaning and preprocessing. (ii) Use examples to illustrate.	10 6	BTL-3	Apply	CO1
8	Discuss the role of EDA in identifying patterns and relationships in data.	16	BTL-3	Apply	CO1
9	Compare and contrast EDA with Classical Statistical Analysis. Provide examples	16	BTL-4	Analyze	CO1
10	Discuss the role of visual aids in EDA. Compare at least five chart types with examples.	16	BTL-4	Analyze	CO1
11	Compare and contrast between supervised and unsupervised EDA approaches.	16	BTL-4	Analyze	CO1
12	Describe the interplay between EDA and hypothesis testing.	16	BTL-4	Analyze	CO1
13	Explain the concept of dimensionality reduction in EDA and describe techniques used.	16	BTL-4	Apply	CO1

14	Formulate merging databases in EDA. Include a real-life use case and explain how keys and joins are managed.	16	BTL-5	Evaluate	CO1
15	Describe the process of identifying and treating outliers using EDA techniques.	16	BTL-5	Evaluate	CO1
16	Explain how software tools like Power BI, Tableau, or Excel support EDA workflows.	16	BTL-6	Create	CO1
17	(i) Using a dataset of monthly sales figures for different product categories, illustrate how you would create a pivot table to summarize the total sales for each product category by quarter. (ii) What are the key components of a pivot table?	10 6	BTL-6	Create	CO1
18	Evaluate the usage of EDA in R vs Python with respect to libraries, speed, and visualization quality.	16	BTL-5	Evaluate	CO1

### UNIT II EDA USING PYTHON

Data Manipulation using Pandas – Pandas Objects – Data Indexing and Selection – Operating on Data – Handling Missing Data – Hierarchical Indexing – Combining datasets – Concat, Append, Merge and Join – Aggregation and grouping – Pivot Tables – Vectorized String Operations.

#### UNIT II PART – A

Q. No	Questions	BT Level	Competence	Course Outcome
1	List the two primary data structures in the pandas library.	BTL-1	Remember	CO2
2	What is the role of the index in a pandas DataFrame?	BTL-1	Remember	CO2
3	Name the three key parameters of the pivot_table() function.	BTL-1	Remember	CO2
4	Mention a command to read a CSV file using pandas.	BTL-1	Remember	CO2
5	What does the head() function do in Pandas?	BTL-1	Remember	CO2
6	Specify the use of the groupby() function.	BTL-1	Remember	CO2
7	What are vectorized string operations in pandas?	BTL-1	Remember	CO2
8	What is the purpose of the drop() method in pandas?	BTL-1	Remember	CO2
9	What happens if you concatenate DataFrames with different indexes?	BTL-1	Remember	CO2
10	Name the function is used to remove duplicate rows in a DataFrame.	BTL-1	Remember	CO2
11	What is the use of the apply() method in pandas?	BTL-1	Remember	CO2
12	What is a ufunc() function and how does it apply to pandas objects?	BTL-1	Remember	CO2
13	What is data aggregation in pandas?	BTL-1	Remember	CO2
14	How do you create a pandas Series from a Python list?	BTL-2	Understand	CO2
15	Differentiate between a pandas Series and a DataFrame.	BTL-2	Understand	CO2
16	How can you create a pandas DataFrame from a dictionary of lists?	BTL-2	Understand	CO2

17	Which method is used to display the top 5 rows of a DataFrame?	BTL-2	Understand	CO2
18	How do you select a single column from a DataFrame by its name?	BTL-2	Understand	CO2
19	How do you determine the dimensions of a DataFrame?	BTL-2	Understand	CO2
20	Differentiate between loc and iloc for data selection in pandas.	BTL-2	Understand	CO2
21	Which method is used to identify missing or null values in a DataFrame?	BTL-2	Understand	CO2
22	Is there any difference between the isnull() and isna() methods in pandas? Justify your answer.	BTL-2	Understand	CO2
23	How would you replace all missing values in a DataFrame with the value 0?	BTL-2	Understand	CO2
24	Write a command to convert a Series of strings to uppercase.	BTL-2	Understand	CO2
25	How does the append() method differ from the concat() function?	BTL-2	Understand	CO2

#### UNIT II PART – B

Q. No	Question	Mark	BT Level	Competence	Course Outcome
1	Explain in detail the different types of pandas objects with suitable examples.	16	BTL-3	Apply	CO2
2	Discuss the different indexing and selection techniques in pandas.	16	BTL-3	Apply	CO2
3	Describe the operations that can be performed on pandas Series and DataFrames.	16	BTL-3	Apply	CO2
4	Explain handling of missing data in pandas. Illustrate with examples.	16	BTL-3	Apply	CO2
5	Explain different types of joins in pandas. Illustrate with examples using merge().	16	BTL-3	Apply	CO2
6	Describe how to combine datasets in pandas using following methods: (i) concat (ii) append (iii) Merge (iv) join.	4 4 4 4	BTL-3	Apply	CO2
7	Explain the use of lambda functions and apply() in pandas.	16	BTL-3	Apply	CO2
8	Describe reshaping and pivoting operations in pandas using (i) stack() (ii) unstack() (iii) melt()	5 5 6	BTL-3	Apply	CO2

9	Compare and contrast concat(), append(), and merge() in pandas. Provide examples.	16	BTL-4	Analyze	CO2
10	Write a pandas program that utilizes the groupby() functionality to analyze a dataset.	16	BTL-4	Analyze	CO2
11	Explain how to reshape data using pivot() and pivot_table() with examples.	16	BTL-4	Analyze	CO2
12	Compare and contrast different pandas methods for merging and joining datasets based on index and columns.	16	BTL-4	Analyze	CO2
13	Explain the purpose of the crosstab() function in pandas. How does it differ from a pivot table?	16	BTL-4	Analyze	CO2
14	Describe the process of importing and exporting data using pandas	16	BTL-5	Evaluate	CO2
15	Describe the "split-apply-combine" strategy that underlies the groupby() operation in pandas.	16	BTL-5	Evaluate	CO2
16	Write a complete pandas program that reads data from a CSV file into a DataFrame. The program should then perform data cleaning by handling any missing values and removing duplicate rows. Compute and display summary statistics for the numerical columns.	16	BTL-6	Create	CO2
17	Write a pandas program to construct a pivot table from a given DataFrame.	16	BTL-6	Create	CO2
18	Demonstrate string operations in pandas using vectorized string methods with examples.	16	BTL-6	Create	CO2

### UNIT – III UNIVARIATE ANALYSIS

Introduction to Single Variable: Distribution Variables – Numerical Summaries of Level and Spread – Scaling and Standardizing – Inequality

Q. No	Questions	BT Level	Competence	Course Outcome
<b>UNIT III PART – A</b>				
1	Define the term mean of a data set.	BTL-1	Remember	CO3
2	What is the median? How is it calculated?	BTL-1	Remember	CO3
3	Give an example of a mode in a data set.	BTL-1	Remember	CO3
4	What does the range measure in a distribution?	BTL-1	Remember	CO3
5	Define InterQuartileRange.	BTL-1	Remember	CO3
6	State the formula for min-max scaling.	BTL-1	Remember	CO3
7	What is meant by a skewed distribution?	BTL-1	Remember	CO3
8	What is a box plot used for?	BTL-1	Remember	CO3
9	If the mean is 50 and the standard deviation is 5, what is the z-score of 60?	BTL-2	Understand	CO3
10	Write the formula for sample standard deviation.	BTL-2	Understand	CO3
11	What is scaling in the context of statistical data?	BTL-1	Remember	CO3

12	Define the variance of a data set.	BTL-1	Remember	CO3
13	Name one real-world use of inequality measurement.	BTL-1	Remember	CO3
14	Differentiate between population and sample mean.	BTL-2	Understand	CO3
15	When is median a better measure of central tendency than mean?	BTL-2	Understand	CO3
16	Why is the denominator for sample variance typically $n-1$ instead of $n$ ?	BTL-2	Understand	CO3
17	Differentiate. between IQR and standard deviation.	BTL-2	Understand	CO3
18	Distinguish between absolute and relative measures of spread.	BTL-2	Understand	CO3
19	Give the formula for the coefficient of variation.	BTL-2	Understand	CO3
20	Define standardizing a variable.	BTL-1	Remember	CO3
21	What is a z-score and what does it represent?	BTL-1	Remember	CO3
22	When is the mode the most appropriate measure of central tendency?	BTL-2	Understand	CO3
23	Give an example of a positively skewed distribution.	BTL-2	Understand	CO3
24	Find the z-score of 120 if the mean is 100 and the standard deviation is 10.	BTL-2	Understand	CO3
25	State the formula for calculating a Z-score.	BTL-1	Remember	CO3

UNIT III PART – B					
Q. No	Question	Mark	BT Level	Competence	Course Outcome
1	Explain the concepts of measures of central tendency. Discuss the properties, advantages, and disadvantages of the mean, median, and mode.	16	BTL-3	Apply	CO3
2	Discuss the concept of skewness and kurtosis in the context of the shape of a distribution. Use diagrams to illustrate symmetric, positively skewed, and negatively skewed distributions.	16	BTL-3	Apply	CO3
3	Explain the importance of data visualization in understanding the distribution of a single variable. Discuss the strengths and weaknesses of at least three different graphical methods for displaying single variable data	16	BTL-3	Apply	CO3
4	Explain the concept of z-scores and their application in real-world data analysis.	16	BTL-3	Apply	CO3
5	Illustrate the calculation of percentiles, quartiles, and IQR using a data set.	16	BTL-3	Apply	CO3
6	Compare and contrast the use of a histogram and a box plot for visualizing a single variable distribution.	16	BTL-4	Analyze	CO3
7	What is the empirical rule (or the 68-95-99.7 rule)? Explain the conditions under which this rule applies. For a normally distributed dataset with a mean of 100 and a standard deviation of 15, what percentage of the data would you expect to fall between 85 and 115? Between 70 and 130? Above 145?	16	BTL-4	Analyze	CO3
8	Explain in detail how to construct a box-and-whisker plot. How can a box plot be used to compare the distributions of a single variable across different groups?	16	BTL-5	Evaluate	CO3

9	Consider the following dataset of exam scores: 65, 72, 78, 80, 82, 85, 85, 88, 90, 95. (i) Calculate the mean, median, and mode. (ii) Calculate the range, variance, and standard deviation. (iii) Calculate the interquartile range (IQR).	6 5 5	BTL-6	Create	CO3
10	Discuss the properties of the mean and standard deviation. Explain how these measures are affected by linear transformations of the data	16	BTL-3	Apply	CO3
11	Discuss the relationship between the concepts of "level" and "spread" in a distribution. Why is it insufficient to only report a measure of central tendency without a corresponding measure of dispersion? Provide an example to illustrate your point.	16	BTL-3	Apply	CO3
12	Suppose you are given the income distribution for two different small towns. Town A has a Gini coefficient of 0.25, and Town B has a Gini coefficient of 0.55. (i) Interpret the meaning of these Gini coefficients. Which town has a more equal income distribution? (ii) Sketch possible Lorenz curves for both towns on the same diagram.	8 8	BTL-3	Apply	CO3
13	Compare and contrast the mean, median, and mode with appropriate graphical illustrations.	16	BTL-4	Analyze	CO3
14	Analyze a skewed dataset using numerical summaries and graphical tools.	16	BTL-4	Analyze	CO3
15	A manufacturing company produces bolts with a target diameter of 10mm. A sample of 10 bolts has the following diameters: 9.8, 10.1, 10.2, 9.9, 10.0, 10.3, 9.7, 10.1, 10.0, 9.9. (i) Calculate the mean and standard deviation of the bolt diameters. (ii) If the production process is standardized, what would be the Z-score for a bolt with a diameter of 10.4mm?	8 8	BTL-4	Apply	CO3
16	A manufacturing company produces bolts with a target diameter of 10mm. A sample of 10 bolts has the following diameters: 9.8, 10.1, 10.2, 9.9, 10.0, 10.3, 9.7, 10.1, 10.0, 9.9. (i) Calculate the mean and standard deviation of the bolt diameters. (ii) The company considers a bolt to be defective if its diameter is more than 0.2mm away from the target. How many defective bolts are in this sample?	8 8	BTL-5	Evaluate	CO3
17	Given a dataset of house prices (in thousands of dollars): 250, 275, 300, 325, 350, 400, 450, 500, 700. (i) Calculate the mean and standard deviation of the house prices. (ii) Standardize the entire dataset by converting each price to a Z-score. (iii) Apply min-max scaling to transform the data into a range of [0, 1].	6 5 5	BTL-6	Create	CO3

18	The following data represents the daily commute time (in minutes) for a group of employees: 35, 45, 50, 55, 60, 60, 65, 70, 75, 120. (i) Calculate the mean and median commute time. (ii) Calculate the standard deviation and the IQR.	8 8	BTL-6	Create	CO3
----	---	--------	-------	--------	-----

UNIT – IV BIVARIATE ANALYSIS				
Relationships between Two Variables – Percentage Tables – Analysis Contingency Tables Handling Several Batches – Scatterplots and Resistant Lines.				
Q. No	Questions	BT Level	Competence	Course Outcome
<b>UNIT IV PART – A</b>				
1	Define a contingency table.	BTL-1	Remember	CO4
2	What does a scatterplot represent?	BTL-1	Remember	CO4
3	What type of data is required for a scatterplot?	BTL-1	Remember	CO4
4	Name two characteristics of a resistant line.	BTL-1	Remember	CO4
5	State one limitation of using percentage tables.	BTL-1	Remember	CO4
6	Define marginal totals in a contingency table.	BTL-1	Remember	CO4
7	Mention one use of a contingency table.	BTL-1	Remember	CO4
8	State one advantage of using resistant lines over traditional lines of best fit.	BTL-1	Remember	CO4
9	What is the role of independent and dependent variables in scatterplots?	BTL-1	Remember	CO4
10	What does a flat line in a scatterplot indicate?	BTL-1	Remember	CO4
11	What is an outlier in a scatterplot?	BTL-1	Remember	CO4
12	Specify the significance of analyzing several batches together.	BTL-1	Remember	CO4
13	What is a proportional relationship?	BTL-1	Remember	CO4
14	Differentiate between row and column percentages in a table.	BTL-2	Understand	CO4
15	How is correlation represented in a scatterplot?	BTL-2	Understand	CO4
16	How do you identify clusters in a scatterplot?	BTL-2	Understand	CO4
17	Differentiate between absolute and relative frequencies.	BTL-2	Understand	CO4
18	Write the importance of using several batches for data analysis.	BTL-2	Understand	CO4
19	How do you find the median of a batch?	BTL-2	Understand	CO4
20	Give an example for negative linear relationship.	BTL-2	Understand	CO4
21	Identify the term “resistant line”.	BTL-2	Understand	CO4
22	Distinguish between a contingency table and a scatterplot.	BTL-2	Understand	CO4
23	How do you interpret the relationship between two categorical variables?	BTL-2	Understand	CO4
24	Give one real-world example of a contingency table.	BTL-2	Understand	CO4
25	How does a scatterplot help in predicting trends?	BTL-2	Understand	CO4

UNIT IV PART – B					
Q. No	Question	Mark	BT Level	Competence	Course Outcome
1	Draw a scatterplot for a given data set and interpret the relationship between the variables.	16	BTL-3	Apply	CO4
2	Explain the steps in converting a frequency table into a percentage table. Provide examples.	16	BTL-3	Apply	CO4
3	Illustrate and explain how resistant lines can be used to interpret non-linear trends.	16	BTL-3	Apply	CO4
4	Given a table of values, calculate both row and column percentages and interpret the results.	16	BTL-3	Apply	CO4
5	Use a real-life example to demonstrate how percentage tables can help in decision-making.	16	BTL-3	Apply	CO4
6	Derive a resistant line from a given dataset, step-by-step, and explain its meaning.	16	BTL-3	Apply	CO4
7	Given a complex dataset, organize it into a contingency table and interpret results using percentages.	16	BTL-3	Apply	CO4
8	Discuss the role of contingency tables in epidemiological studies. Provide sample data.	16	BTL-3	Apply	CO4
9	Compare and contrast the use of scatterplots and percentage tables in identifying relationships.	16	BTL-4	Analyze	CO4
10	Analyze a multi-batch dataset using resistant lines. Explain how the line is determined and its significance.	16	BTL-4	Analyze	CO4
11	Discuss the limitations and advantages of analyzing several batches simultaneously.	16	BTL-4	Analyze	CO4
12	Explain how to use conditional percentages to analyze relationships in a table.	16	BTL-4	Analyze	CO4
13	Demonstrate how to construct and analyze a scatterplot with three batches of data.	16	BTL-4	Apply	CO4
14	Describe the process of identifying clusters and outliers in a scatterplot. Provide examples.	16	BTL-5	Evaluate	CO4
15	Discuss how outliers affect the line of best fit and resistant line. Use diagrams to illustrate.	16	BTL-5	Evaluate	CO4
16	Construct a percentage contingency table from raw data and comment on any apparent associations.	16	BTL-6	Create	CO4
17	Construct and analyze a contingency table for a given set of categorical data.	16	BTL-6	Create	CO4
18	Explain the steps for constructing a resistant line and compare with a least-squares regression line.	16	BTL-4	Analyze	CO4

**UNIT – V MULTIVARIATE AND TIME SERIES ANALYSIS**

Introducing a Third Variable – Causal Explanations – Three-Variable Contingency Tables and Beyond – Fundamentals of TSA – Characteristics of time series data – Data Cleaning – Time-based indexing – Visualizing – Grouping – Resampling.

Q. No	Questions	BT Level	Competence	Course Outcome
<b>UNIT V PART – A</b>				
1	What is a third variable in data analysis?	BTL-1	Remember	CO5
2	Define a confounding variable.	BTL-1	Remember	CO5
3	What does causality imply in statistical terms?	BTL-1	Remember	CO5
4	What is meant by time series data?	BTL-1	Remember	CO5
5	List two key features of time series data.	BTL-1	Remember	CO5
6	Define time-based indexing.	BTL-1	Remember	CO5
7	What is a seasonal pattern in time series?	BTL-1	Remember	CO5
8	What does resampling mean in time series analysis?	BTL-1	Remember	CO5
9	Mention one use of grouping in time series.	BTL-1	Remember	CO5
10	State one advantage of a line chart in time series visualization.	BTL-1	Remember	CO5
11	State one method of handling missing time-based data.	BTL-1	Remember	CO5
12	What kind of chart is best for showing seasonality?	BTL-1	Remember	CO5
13	What is a temporal granularity?	BTL-1	Remember	CO5
14	Write the purpose of a three-variable contingency table.	BTL-2	Understand	CO5
15	Differentiate between association and causation.	BTL-2	Understand	CO5
16	Give one example of a time-based variable.	BTL-2	Understand	CO5
17	Why is missing data a problem in time series?	BTL-2	Understand	CO5
18	How is a time series different from cross-sectional data?	BTL-2	Understand	CO5
19	Give one reason to visualize time series data.	BTL-2	Understand	CO5
20	Give an example of a situation requiring a three-way contingency table.	BTL-2	Understand	CO5
21	How does a lurking variable affect analysis?	BTL-2	Understand	CO5
22	Write one Python library used for time series visualization.	BTL-2	Understand	CO5
23	Identify the role of forecasting in TSA.	BTL-2	Understand	CO5
24	Write the characteristics of time series data.	BTL-2	Understand	CO5
25	Give an example of a cyclic trend.	BTL-2	Understand	CO5

**UNIT V PART – B**

Q. No	Question	Mark	BT Level	Competence	Course Outcome
1	Explain the role of a third variable in determining whether an observed relationship is causal or spurious.	16	BTL-3	Apply	CO5
2	Describe in detail the key components of time series data.	16	BTL-3	Apply	CO5
3	Illustrate with examples how data cleaning improves the quality of time series analysis.	16	BTL-3	Apply	CO5
4	Explain the process of time-based indexing and its significance in time series analysis.	16	BTL-3	Apply	CO5

5	Explain the process and purpose of resampling in time series. Use code and visualizations if needed.	16	BTL-3	Apply	CO5
6	Demonstrate how a third variable can alter the interpretation of the relationship between two other variables.	16	BTL-3	Apply	CO5
7	Explain different methods for imputing missing time series data	16	BTL-3	Apply	CO5
8	Discuss the importance of choosing the right time granularity in time-based data analysis.	16	BTL-3	Apply	CO5
9	Compare upsampling and downsampling with practical examples.	16	BTL-4	Analyze	CO5
10	Explain the use of ARIMA models in TSA and their applications.	16	BTL-4	Analyze	CO5
11	Analyze ethical considerations in causal explanations when working with three variables.	16	BTL-4	Analyze	CO5
12	Compare different time series resampling strategies and their impact on trend analysis.	16	BTL-4	Analyze	CO5
13	Design and interpret a multivariate time series visualization.	16	BTL-4	Analyze	CO5
14	Formulate how to prepare a raw CSV dataset with time data for time series analysis in Python.	16	BTL-5	Evaluate	CO5
15	Describe how to detect and handle missing values in time series data using pandas.	16	BTL-5	Evaluate	CO5
16	Construct a time series plot using real or simulated data and explain how it informs business or scientific decisions.	16	BTL-6	Create	CO5
17	Construct and interpret a three-variable contingency table. Identify possible interactions.	16	BTL-6	Create	CO5
18	Discuss the importance of evaluating the performance of a time series model.	16	BTL-5	Evaluate	CO5